

Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Ch6>
 RPL-83M4 is from the library RPL-1 constructed by the group of
 Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2
 IMPORTANT: This sequence is not the entire insert of clone RPL-83M4
 It may be shorter because we sequence overlapping sections only
 once, except for a 100 base overlap.
 The true left end of clone RPL-83M4 is at 1 in this sequence. The
 true left end of clone RPL-91J24 is at 104129 in this sequence.

FEATURES

Source
 1..104228
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RPL-83M4"
 /clone_lib="RPL-1"
 1302..1605
 repeat_region repeat: matches 1..303 of consensus"
 2223..2533
 repeat_region repeat: matches 1866..2190 of consensus"
 2573..2617
 repeat_region repeat: matches 202..247 of consensus"
 2581..2634
 repeat_region repeat: matches 2693..2747 of consensus"
 2637..2861
 repeat_region repeat: matches 1..262 of consensus"
 3966..4024
 repeat_region repeat: matches 35..77 of consensus"
 4613..4720
 repeat_region repeat: matches 1..278 of consensus"
 4759..5035
 repeat_region repeat: matches 1..150 of consensus"
 5613..5697
 repeat_region repeat: matches 65..150 of consensus"
 5746..6018
 repeat_region repeat: matches 169..450 of consensus"
 6681..6805
 repeat_region repeat: matches 61..185 of consensus"
 7197..7476
 repeat_region repeat: matches 140 mer 85% conserved"
 7544..7704
 repeat_region repeat: matches 2355..2530 of consensus"
 7835..7962
 repeat_region repeat: matches 2609..2746 of consensus"
 8707..8977
 repeat_region repeat: matches 1..268 of consensus"
 10041..10339
 repeat_region repeat: matches 1..301 of consensus"
 11963..12086
 repeat_region repeat: matches 1..124 of consensus"
 12102..12143
 repeat_region repeat: matches 98..142 of consensus"
 12144..12456
 repeat_region repeat: matches 1..304 of consensus"
 12457..12617
 repeat_region repeat: matches 142..282 of consensus"
 12648..12786
 repeat_region repeat: matches 2596..2745 of consensus"
 13356..13668
 repeat_region repeat: matches 1..310 of consensus"
 13845..14157
 repeat_region repeat: matches 1..313 of consensus"
 14484..14623
 repeat_region repeat: matches 70..212 of consensus"
 15188..15497
 repeat_region repeat: matches 1..311 of consensus"
 15498..15557
 repeat_region repeat: matches 1..311 of consensus"
 15931..16276
 repeat_region repeat: matches 2383..2731 of
 consensus"

repeat_region repeat: matches 6..312 of consensus"
 repeat_region repeat: matches 627..2383 of
 consensus"
 misc-feature
 16277..16587
 /note="AluB repeat: matches 6..312 of consensus"
 16588..18382
 /note="Rigge4(Zomb1) repeat: matches 627..2383 of
 consensus"
 16997..17123
 /note="Sequence from overlapping clone B3B1 (AL512354).
 Assembly confirmed by restriction digest."
 18383..18643
 /note="AluB repeat: matches 7..275 of consensus"
 18644..19263
 /note="Rigge4(Zomb1) repeat: matches 1..627 of consensus"
 19527..19632
 /note="MIR repeat: matches 154..258 of consensus"
 20848..20937
 /note="MIR repeat: matches 146..240 of consensus"
 20940..21242
 /note="MIR repeat: matches 70..429 of consensus"
 22808..23017
 /note="MIR repeat: matches 70..284 of consensus"
 23035..23336
 /note="AluB repeat: matches 1..309 of consensus"
 23523..23658
 /note="AluB repeat: matches 1..139 of consensus"
 23659..23917
 /note="AluB repeat: matches 21..285 of consensus"
 24084..24278
 /note="AluB repeat: matches 1..195 of consensus"
 25005..25087
 /note="MIR repeat: matches 92..185 of consensus"
 26139..26440
 /note="AluB repeat: matches 1..299 of consensus"
 26889..27041
 /note="MIR repeat: matches 78..252 of consensus"
 27623..28017
 /note="MIR repeat: matches 1..364 of consensus"
 29032..29383
 /note="MIR repeat: matches 1..354 of consensus"
 29985..30289
 /note="MIR repeat: matches 1..307 of consensus"
 30313..30457
 /note="MIR repeat: matches 2..189 of consensus"
 32411..33100
 /note="MIR repeat: matches 5426..6134 of consensus"
 33301..33442
 /note="MIR repeat: matches 71..232 of consensus"
 34707..34752
 /note="MIR repeat: matches 23 copies 2 mer 76% conserved"
 34819..34990
 /note="MIR repeat: matches 5972..6143 of consensus"
 34992..35031
 /note="MIR repeat: matches 10 copies 4 mer 85% conserved"
 35362..36232
 /note="CpG Island"
 evidence-not-experimental
 37387..37697
 /note="AluB repeat: matches 1..311 of consensus"
 37710..37993
 /note="AluB repeat: matches 1..286 of consensus"
 38849..39089
 /note="MIR repeat: matches 2..262 of consensus"
 40658..40945
 /note="AluB repeat: matches 1..288 of consensus"
 41221..41256
 /note="MIR repeat: matches 1806..1889 of consensus"
 41267..41347
 /note="MIR repeat: matches 1806..1889 of consensus"
 41323..41464
 /note="MIR repeat: matches 2007..2143 of consensus"
 41487..41557
 /note="MIR repeat: matches 6099..6167 of consensus"
 41558..41965
 /note="MIR repeat: matches 1..426 of consensus"

/db_xref="GI:3248918"
 /translation="MKDRLAELLDLSKYDQDPFGDDFEDFSPHEDIVETPHILESL
 YRIDRIDENOLIVADVRLKONARFLTSMRLSSIKRDNSTAKARAGEVTHC
 KLRAMELSEAAOHPHSAVARISRAOYNALTLTFORAMHDYNOAEKRONCKIR
 IOROLEIMKEVSGDOIEMFEGKMDVSENLADYKRGCPPTSRRAATNCCMR
 AAIKRVHELFLOMAVLEKADTLNVLNVTYDITGQAKQVAKAYEKNPCR
 TLCCFCCPLK"
 BASE COUNT 281 a 409 c 399 g 191 t
 ORIGIN

Query Match 44.1%; Score 570.6; DB 9; Length 1280;
 Best Local Similarity 98.9%; Pred. No. 1.6e-79;
 Matches 606; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

OY 1 acccagcgccagctgagatcatatggaagaagatctcgcgccagccagatcgaagacatg 60
 |||||||
 DB 664 ATCCAGCGCCAGCTGAGATCATGAGGCAAGAGTCTCGGCGCACAGATCGAGACATG 723
 OY 61 ttcgagcagggtgaagtggagcgtgtttccgagaacttgctgacgacgtgaaggcgcg 120
 |||||||
 DB 724 TTCCGAGCAGGTAAGTGGACGCTTTCCGAGAACTTGCTGGCGCAGTGAAGGCG -CG 782
 OY 121 cggcgccgctcgaagagtcgagagccgcccagcgaactgtgctgagcctggagagcg - 179
 |||||||
 DB 783 CGGCGCGCCACAGACGATCGAGAGCCCGCACCGCAACTGCTGCGCTGAGAGCGCG 842
 OY 180 catccgagcagctacgaagatctctctgagatgagcgtgctggtggaagaagcagcgca 239
 |||||||
 DB 843 CATCCGCGAGATACAGACACTCTCTTTCAGATGCGGCTGCTGAGAGAGAGCGCGA 902
 OY 240 caacctgaagcgtcatcgaagtcacgaagacggtgactacacggcgagcgca 239
 |||||||
 DB 903 CACCTGAAAGTATCATGAGCTCAACGTACAAAAGACGATGACTACACCGCGCAGCGCAA 962
 OY 300 ggcgagagtgaggaagcggctgcaatgagagaggaagaccctgcccgaagcctctgctg 359
 |||||||
 DB 963 GCGCAGAGTGGAGAGCGGTGACGAGAGAGAACCCCTGCGGAGCCTCTGCTG 1022
 OY 360 ctctcgtctgctcctgctcgaagtagagcagcgcccgcgccgacacgcgcccacatccagac 419
 |||||||
 DB 1023 CTTTGTGCTGCTGCTGCTCAAGTACAGGCGCGCGCGCGCGCGCCATCCAGAC 1082
 OY 420 catgagcgcgctgggaagagcgtcaacaaagccggagagctctgctcagaggaagttgc 479
 |||||||
 DB 1083 CATGAGAGCGCGTGGGAAAGACG -CACCAAGCGCGGAGCTCTGCCCTGAGGAGATTGC 1141
 OY 480 cccaacccttcggaaccagctctttagaagaagaagcgcaggttcaagaattgcaaac 539
 |||||||
 DB 1142 CCCAACCCCTTCCGGAATCAGCTCTTACAAAAGAAAGCCAGGTTCAAGAAATTGCAAC 1201
 OY 540 cagcgtctgctggaagaatggttaattgatacgcgtcgaatgattcttcagataagtag 599
 |||||||
 DB 1202 CAGCTGTGCTTGGAAAGATGTTACTTGATACCCGTCATGATTTCTTCAAGTAAAGTAG 1261
 OY 600 attcccaacaagt 612
 |||||||
 DB 1262 ATTCCACACTCGT 1274

RESULT 4
 AF071504 1000 bp mRNA linear PRI 22-JUN-1998
 LOCUS AF071504
 DEFINITION Homo sapiens syntaxin 11 mRNA, complete cds.
 ACCESSION AF071504
 VERSION AF071504.1 GI:3243239
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1000)
 AUTHORS Valdez,A.C., Cabanillos,J.-P., Brown,M.J. and Roche,P.A.

TITLE Syntaxin 11 is a novel SNARE protein associated with the
 Trans-Golgi network and late endosomes
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1000)
 AUTHORS Valdez A.C., Cabanillos J.-P., Brown, M.J. and Roche, P.A.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUN-1998) NCI/EIB, NIH, 10 Center Drive, Bethesda, MD
 20892, USA

FEATURES
 source
 1..1000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /issue_type="placenta"
 55..918
 /note="SNARE protein"
 /codon_start=1
 /product="syntaxin 11"
 /protein_id="AAC24004.1"
 /db_xref="GI:3243240"
 /translation="MKDRLAELLDLSKYDQDPFGDDFEDFSPHEDIVETPHILESL
 YRIDRIDENOLIVADVRLKONARFLTSMRLSSIKRDNSTAKARAGEVTHC
 KLRAMELSEAAOHPHSAVARISRAOYNALTLTFORAMHDYNOAEKRONCKIR
 IOROLEIMKEVSGDOIEMFEGKMDVSENLADYKRGCPPTSRRAATNCCMR
 AAIKRVHELFLOMAVLEKADTLNVLNVTYDITGQAKQVAKAYEKNPCR
 TLCCFCCPLK"

BASE COUNT 224 a 327 c 307 g 142 t
 ORIGIN

Query Match 35.0%; Score 453.4; DB 9; Length 1000;
 Best Local Similarity 99.0%; Pred. No. 2.9e-61;
 Matches 465; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 atccagcgccagctgagatcatatggaagaagtcgcggcgccagccagatcgaagacatg 60
 |||||||
 DB 535 ATCCAGCGCCAGCTGAGATCATGAGGCAAGATCTCGGCGCACAGATCGAGACATG 594
 OY 61 ttcgagcagggtgaagtggagcgtgtttccgagaacttgctgacgacgtgaaggcgcg 120
 |||||||
 DB 595 TTCCGAGCAGGTAAGTGGAGACGCTTTCCGAGAACTTGCTGCGCGCAGTGAAGGCGCG 654
 OY 121 cggcgccgctcgaagagatcgaagagccgacacgcgcaactgctgctggagagcgcg 180
 |||||||
 DB 121 GCGCAGAGTGGAGAGCGGTGACGAGTCTTCCGAGAACTTGCTGCGCGCAGTGAAGGCGCG 654
 OY 181 atccgagcagctacgaagcgtctcttcgagatgagcgtgctggtggaagaagcagcgac 240
 |||||||
 DB 715 ATCCGCGACGTACACGAGCTCTTCTTCAAGTGGCGGTGCTGAGGAAAGACGCGGAC 774
 OY 241 accctgaagctcatcgaagctcaacgtaacaaagacggtcgaactacacggcgccagcgcaag 300
 |||||||
 DB 775 ACCCTGAAGCTCATGCGATCAACGATCAAAAAGAGCGGTGACGATACACGGCGCAGCGCAAG 834
 OY 301 ggcgaggtgaggaagcgcgtgagtaagagagaagaacccctgcggagccctctgctgc 360
 |||||||
 DB 835 GCGCAGAGTGGAGAGCGGTGACGATGAGAGAGAACCCCTGCGCAGCTCTGCTGC 894
 OY 361 ttcgctgctcctgctcgaagtagcagcgccggcgccgacacggccatcccaaac 420
 |||||||
 DB 895 TTCTGCTGCTGCTGCTCAAGTACAGGCGCGCGCGCGCGCCACCCCATCCCAAC 954
 OY 421 atgagcgcgctggaagaagacgtcaacaaagccggagcgtgctgctc 467
 |||||||
 DB 955 ATGAGCGCGCTGGAAGAGACG -CACCAAGCGCGGAGCTCTGCGCT 1000

RESULT 5
 AC091681 207379 bp DNA linear HMG 29-JAN-2002
 LOCUS AC091681
 DEFINITION Mus musculus chromosome 10 clone rp23-111d4 strain C57BL/6J,
 WORKING DRAFT SEQUENCE, 29 unordered pieces.
 ACCESSION AC091681
 VERSION AC091681.27 GI:18390231

Accession	Gene	Species	Protein	Length	Score	E-value	Identity	Positives	Gaps	Conserved Domains
U00096	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00097	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00098	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00099	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00100	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00101	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00102	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00103	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00104	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00105	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00106	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00107	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00108	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00109	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00110	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00111	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00112	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00113	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00114	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00115	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00116	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00117	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00118	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00119	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00120	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00121	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00122	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00123	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00124	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00125	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	5S rRNA
U00126	16S rRNA	<i>Escherichia coli</i>	16S rRNA	1641	100.0	0.0	100	1641	0	16S rRNA
U00127	23S rRNA	<i>Escherichia coli</i>	23S rRNA	2324	100.0	0.0	100	2324	0	23S rRNA
U00128	5S rRNA	<i>Escherichia coli</i>	5S rRNA	120	100.0	0.0	100	120	0	

[illegible]

|||||
Db 448 ATCCAGCGCAGCTAGAGATCACTGGAGAGACCACCACTGACGAGGAGCGTGGAGGAGATG 507
OY 61 ttccgagcaggttaagtggagcgtgtttccgagaacttgctggcagcagctgaagggcgcg 120
Db 508 CTGGAGAGCGGGAAGCGCCATCTTCATCTCGGACATATATACATTCACAGATTACT 567
OY 121 cggagcgccttaacagagatcgagagcgccacgcgaaactgctgcgccttgagagcgcg 180
Db 568 AGGCAAGCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
OY 181 atccgagcagctacagagcttcttccgagagtggtgctggttgagagagcgccgac 240
Db 628 ATCCGAGAGCTCAGGAGATGTTATGATGATGATGATGATGATGATGATGATGATG 687
OY 241 accctgaagctacgagctacgagctacgagctacgagctacgagctacgagctacgag 300
Db 688 ATGCTCAACACATTCAGAGAAACGTGGTAAGTCCGTGATTCAGTGGAGCAGCCCAAG 747
OY 301 gcgcaggtgcgagagcgctgcagctacgagagagagagagagagagagagagagag 360
Db 748 GAAGAGACTAAGAAAGCCATCAATACAGAGCAAGCCAGAGCGGGGTGTGCTGTGCT 807
OY 361 ttctgttcctctgc 375
Db 808 CTCGGCCAGCAGCTGC 822

RESULT 10
AR111444
LOCUS AR111444 486 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 13 from patent US 6127149.
ACCESSION AR111444
VERSION AR111444.1 GI:12828292
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 486)
AUTHORS Hirai, Y., Koshida, S. and Oka, Y.
TITLE Modified epimorphin
JOURNAL Patent: US 6127149-A 13-03-OCT-2000;
FEATURES
Source location/Qualifiers
1..486
BASE COUNT 152 a 110 c 139 g 85 t
ORIGIN

Query Match 9.6%; Score 124.8; DB 6; Length 486;
Best Local Similarity 60.7%; Pred. No. 3.9e-10;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

OY 1 atccgagcagctgagatcatgagcagagagagctgcggcgagcagagatg 60
Db 133 ATCCAGCGCAGCTAGAGATCACTGGAGAGACCACCACTGACGAGGAGCTGGAAGATG 192
OY 61 ttccgagcaggttaagtggagcgtgtttccgagaacttgctggcagcagctgaagggcgcg 120
Db 193 CTGGAGAGCGGGAAGCGCCATCTTCATCTCGGACATATATACATTCACAAATCACT 252
OY 121 cggagcgccttaacagagatcgagagcgccacgcgaaactgctgcgccttgagagcgcg 180
Db 253 AGGCAAGCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
OY 181 atccgagcagctacagagcttcttccgagagtggtgctggttgagagagcgccgac 240
Db 313 ATCCGAGAGCTCAGGAGATGTTATGATGATGATGATGATGATGATGATGATGATG 372
OY 241 accctgaagctacgagctacgagctacgagctacgagctacgagctacgagctacgag 300
Db 373 ATGCTCAACACATTCAGAGAAACGTGGTAAGTCCGTGATTCAGTGGAGCAGCCCAAG 432

OY 301 gcgcaggtgcgagagcgctgcagctacgagagagag 336
Db 433 GAAGAGACGAGAGAAAGCCATCAATACAGAGCAAG 468

RESULT 11
E12728
LOCUS E12728 486 bp DNA linear PAT 24-JUN-1998
DEFINITION DNA encoding mouse Epimorphin-derived peptide.
E12728
ACCESSION E12728
VERSION E12728.1 GI:3251560
KEYWORDS JP 1997065885-A/7.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 486)
AUTHORS Koshida, S., Oka, Y. and Hirai, Y.
TITLE TAILORED DERIVATIVE OF EPIMORPHIN
JOURNAL Patent: JP 1997065885-A 7 11-MAR-1997;
SUMITOMO ELECTRIC IND LTD
COMMENT
OC None
OS Artificial sequences.
PN JP 1997065885-A/7
PD 11-MAR-1997
PF 29-MAR-1996 JP 1996099684
PR 31-MAR-1995 JP 95P 99960, 19-JUN-1995 JP 95P 175540 PI
PC KOSHIDA SHOHO, OKA YUMIKO, HIRAI YOHEI
PC C12N15/09, C07H21/04, C07K7/06, C07K14/485, PC
C12P21/02//A61K38/00
PC (C12P21/02, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key location/Qualifiers
FT source 1..486
FT mat_peptide 1..483
FT (23) /product='Epimorphin fragment designated FT
aa of mouse
Epimorphin'.
FEATURES
Source location/Qualifiers
1..486
BASE COUNT 152 a 110 c 139 g 85 t
ORIGIN

Query Match 9.6%; Score 124.8; DB 6; Length 486;
Best Local Similarity 60.7%; Pred. No. 3.9e-10;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

OY 1 atccgagcagctgagatcatgagcagagagagctgcggcgagcagagatg 60
Db 133 ATCCAGCGCAGCTAGAGATCACTGGAGAGACCACCACTGACGAGGAGCTGGAAGATG 192
OY 61 ttccgagcaggttaagtggagcgtgtttccgagaacttgctggcagcagctgaagggcgcg 120
Db 193 CTGGAGAGCGGGAAGCGCCATCTTCATCTCGGACATATATACATTCACAAATCACT 252
OY 121 cggagcgccttaacagagatcgagagcgccacgcgaaactgctgcgccttgagagcgcg 180
Db 253 AGGCAAGCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
OY 181 atccgagcagctacagagcttcttccgagagtggtgctggttgagagagcgccgac 240
Db 313 ATCCGAGAGCTCAGGAGATGTTATGATGATGATGATGATGATGATGATGATGATG 372
OY 241 accctgaagctacgagctacgagctacgagctacgagctacgagctacgagctacgag 300
Db 373 ATGCTCAACACATTCAGAGAAACGTGGTAAGTCCGTGATTCAGTGGAGCAGCCCAAG 432

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 14:13:36 ; Search time 237.5 Seconds

(without alignments)
9361.702 Million cell updates/sec

Title: us-09-762-249-13

Perfect score: 1295
Sequence: 1 atccagcgcagctgagat.....aaaaaaaaaaaaaaaaaaaa 1295

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_032802:*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209.2	16.2	240	AAZ61486	CDNA sequence of t
2	124.8	9.6	486	AAT16088	Mouse epimorphin f
3	124.8	9.6	486	AAT62413	Mouse epimorphin f
4	124.8	9.6	564	AAT16087	Mouse epimorphin f
5	124.8	9.6	564	AAT62412	Mouse epimorphin f
6	124.8	9.6	711	AAT16086	Mouse epimorphin f
7	124.8	9.6	711	AAT62411	Mouse epimorphin f
8	124.8	9.6	798	AAT16090	Mouse epimorphin f
9	124.8	9.6	798	AAT62414	Mouse epimorphin f

10	124.8	9.6	840	AAQ41594	Mouse epimorphin 1
11	124.8	9.6	840	AAQ75249	Mouse epimorphine
12	124.8	9.6	867	AAQ41593	Mouse epimorphin 1
13	124.8	9.6	867	AAQ75248	Mouse epimorphine
14	124.8	9.6	870	AAQ41592	Mouse epimorphin c
15	124.8	9.6	870	AAQ75247	Wild type mouse ep
16	124.8	9.6	2940	AAQ41595	Mouse epimorphin g
17	124.8	9.6	2940	AAQ75250	DNA fragment conta
18	121.6	9.4	911	AAV01548	Rat syntaxin 2 gen
19	118.4	9.1	2088	AAV00371	CDNA encoding SNAR
20	117.2	9.1	2097	AAV01546	Rat syntaxin 1A ge
21	107.2	8.3	1200	AAQ54355	Rat post-synaptic
22	107.2	8.3	3000	AAV01547	Rat syntaxin 1B ge
23	96.6	7.5	486	AAA89595	Mouse syntaxin4 nu
24	96.6	7.5	486	AAA89596	Mouse syntaxin4 nu
25	96.6	7.5	897	AB199515	Mouse ischemic co
26	96.6	7.5	1219	AAA89551	Mouse syntaxin4 nu
27	94.4	7.3	486	AAT16085	Human epimorphin f
28	94.4	7.3	486	AAT62409	Human epimorphin f
29	94.4	7.3	564	AAT16084	Human epimorphin f
30	94.4	7.3	564	AAT62408	Human epimorphin f
31	94.4	7.3	711	AAT16083	Human epimorphin f
32	94.4	7.3	711	AAT62407	Human epimorphin f
33	94.4	7.3	795	AAT16089	Human epimorphin t
34	94.4	7.3	795	AAT62410	Human epimorphin c
35	94.4	7.3	834	AAQ75246	Human epimorphine
36	94.4	7.3	864	AAQ41590	Human epimorphin 1
37	94.4	7.3	864	AAQ75245	Human epimorphine
38	94.4	7.3	867	AAQ41589	Human epimorphin c
39	94.4	7.3	867	AAQ75244	Wild type human ep
40	92.8	7.2	834	AAQ41581	Human epimorphin 1
41	90.6	7.0	892	ABA03884	Human POLY15 nucle
42	90.6	7.0	892	ABA03885	Human POLY16 nucle
43	90.6	7.0	893	ABA03883	Human POLY14 nucle
44	90.6	7.0	1233	AAV01549	Human diagnostic a
45	89.6	6.9	1054	AAV01549	Rat syntaxin 3 gen

ALIGNMENTS

RESULT 1	
ID AAZ61486	standard; CDNA; 240 BP.
XX	
AC AAZ61486;	
XX	
DT 19-JUN-2000	(first entry)
XX	
DE	CDNA sequence of tumour suppressor activated pathway gene TSAP21.
XX	
XX	Tumour suppressor activated pathway gene; TSAP; apoptosis;
KW	tumour suppression; tumour suppressor inhibited pathway gene; TSIP3;
KW	cancer; antiviral; ss.
XX	
OS	Homo sapiens.
XX	
PN	FR2782085-A1.
XX	
PD	11-FEB-2000.
XX	
PF	05-AUG-1998; 98FR-0010077.
XX	
PR	05-AUG-1998; 98FR-0010077.
XX	
PA	(DAUS-) FOND DAUSSET-CEPH JEAN.
XX	
PI	Amson R, Telerman A;
XX	
DR	WPI: 2000-208788/19.
XX	
PT	Nucleic acid sequences useful for gene therapy of cancer and viral
	infections correspond to genes whose expression is induced or inhibited

```

XX JP09065885-A.
XX 11-MAR-1997.
XX 29-MAR-1996; 96JP-0099684.
XX 19-JUN-1995; 95JP-0175540.
XX 31-MAR-1995; 95JP-0099980.
XX (SUME ) SUMITOMO ELECTRIC IND CO.
XX WPI: 1997-220419/20.
XX P-PSDB: AAM14262.
XX Modified epimorphin and related DNA - useful e.g. for treatment of
XX tissues or in artificial organs, or as an ingredient in cosmetics
XX
XX Example 1; Page 13-14; 18pp; Japanese.
XX
XX The invention relates to novel human (AAM14257-9) or mouse (AAM14260-2)
XX epimorphin proteins with replacements, deletions or substitutions in the
XX amino acid sequence. The new epimorphin protein consists of: (a) an
XX N-terminal coiled-coil region; (b) a functional domain in the middle;
XX and (c) a C-terminal coiled-coil region. A hydrophobic region in the
XX C-terminal has been deleted and at least some amino acids have been
XX deleted from the terminals of coiled coil regions (a) and/or (c). This
XX sequence encodes the mouse epimorphin protein lacking the N-terminal
XX 103 amino acids, i.e. the complete N-terminal coiled-coil domain.
XX Epimorphin is a protein which regulates morphogenesis of epithelial
XX tissues: It can be used for treatment of tissues or used directly in
XX artificial organs or as an ingredient in cosmetics, hair tonic, etc.
XX
XX Sequence 486 BP; 152 A; 110 C; 139 G; 85 T; 0 other:
XX
XX Query Match 9.6%; Score 124.8; DB 18; Length 486;
XX Best Local Similarity 60.7%; Pred. No. 3,9e-13;
XX Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
XX
OY 1 atccagcgccagctgagatcagtcaggaagtcgtcgagccagcagatcgagagatg 60
DB 133 atccagcgccagctgagatcagtcaggaagtcgtcgagccagcagcagtcgagagatg 192
OY 61 ttcgagcaggtgtaagtcgagtcgttttcgagaactgtgagcgagtcgagtcgagtcg 120
DB 193 ctgagagagcggaagcgcgtccatcttcacatctcgatataatataacaaatcact 252
OY 121 cgggcccgcctcaacagatcagagagcgcacccggaactgtcgagtcgagtcgagtcg 180
DB 253 agcgaagctcctcaatgagtcgagtcgcgcgcacaaagacatcatgagtcgagagcagc 312
OY 181 atccgagcgctacagagctcttcgagatgagtcgtgtaggaagcagcgagc 240
DB 313 atccgagagctgacagagctgttcacatgagatgagtcgttcgagatcagtcgagtcg 372
OY 241 accctgaagctacagatcagatcagacagtcacaaagagcagtcgagtcgagtcgagtcg 300
DB 373 atgctcaacaacatcagaggaatgtgtgtaactctgtagatcagtcgtgaaactgcagc 422
OY 301 gctcaggtgctgagagcgctgtagtcagtcagtcgagtcgagtcgagtcgagtcgagtc 336
DB 433 gaagagacgaagaagcctacatcaatcacagagcagc 468
XX
XX RESULT 4
XX ID AAT16087 standard: cDNA; 564 BP.
XX AC AAT16087;
XX XX
XX 30-MAY-1996 (first entry)
XX

```

```

DE Mouse epimorphin fragment (3M).
XX
XX Epimorphin; human; mouse; wound; burn; epithelial tissue;
XX diagnosis; treatment; morphogenetic abnormality; cosmetic;
XX hair growth stimulator; ds.
XX
XX Mus musculus.
XX
XX EP698666-A2.
XX
XX 28-FEB-1996.
XX
XX 20-JUN-1995; 95EP-0304270.
XX
XX 31-MAR-1995; 95JP-0099980.
XX 21-JUN-1994; 94JP-0162874.
XX 31-MAR-1995; 95JP-0099979.
XX
XX (SUME ) SUMITOMO ELECTRIC IND CO.
XX
XX Hirai Y, Koshida S, Oka Y;
XX
XX WPI: 1996-118213/13.
XX P-PSDB: AAR92047.
XX
XX Novel polypeptide containing an epimorphin functional domain - has
XX possible benefits in epithelial tissue treatments, e.g. burns and
XX for artificial organs
XX
XX Claim 27; Page 31; 62pp; English.
XX
XX New polypeptides contain a first portion of 5-99 amino acids joined
XX to a second portion contg. at least a functional domain of
XX epimorphin. The first portion may be selected from the peptides
XX given in AAR92029 to AAR92036. The second portion may be full-length
XX epimorphin (see AAR92037 to AAR92042 for human and mouse epimorphins).
XX Fragments (123), (2M), (3M) and (23) of epimorphin given in AAT16083
XX to AAT16090 are used in the prodn. of modified epimorphins.
XX 123: N-terminus to right before C-terminal hydrophobic domain.
XX 2M: amino acid 30 to right before C-terminal hydrophobic domain.
XX 3M: amino acid 105 to right before C-terminal hydrophobic domain.
XX 23: amino acid 105 to right before C-terminal hydrophobic domain.
XX The modified epimorphins are useful for the development of diagnosis
XX and treatment of morphogenetic abnormalities of epithelial tissue or
XX novel remedies for wounds, eg burns, after surgery and for artificial
XX organs. They may also be used as components of cosmetics, hair
XX growth stimulators, etc.
XX
XX Sequence 564 BP; 181 A; 124 C; 162 G; 97 T; 0 other:
XX
XX Query Match 9.6%; Score 124.8; DB 17; Length 564;
XX Best Local Similarity 60.7%; Pred. No. 4e-13;
XX Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
XX
OY 1 atccagcgccagctgagatcagtcggaagtcgtcgagccagcagcagtcgagagatg 60
DB 211 atccagcgccagctgagatcagtcggaagtcgtcgagccagcagcagtcgagagatg 270
OY 61 ttcgagcaggtgtaagtcgagtcgttttcgagaactgtgagcgagtcgagtcgagtcg 120
DB 271 ctgagagagcggaagcgcgtccatcttcacatctcgatataatataacaaatcact 330
OY 121 cgggcccgcctcaacagatcagagagcgcacccggaactgtcgagtcgagtcgagtcg 180
DB 331 agcgaagctcctcaatgagtcgagtcgcgcgcacaaagacatcatgagtcgagagcagc 390
OY 181 atccgagcgctacagagctcttcgagatgagtcgtgtaggaagcagcgagc 240
DB 391 atccgagagctgacagagctgttcacatgagatgagtcgttcgagatcagtcgagtcg 450
OY 241 accctgaagctacagatcagatcagacagtcacaaagagcagtcgagtcgagtcgagtcg 300

```


Query Match 9.6%; Score 124.8; DB 17; Length 711;
Best Local Similarity 60.7%; Pred. No. 4.2e-13;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1 atccagcgcaagcttgagatcatatggcaagaagctctcgagaccacatcgagacatg 60
|||||
DB 358 atccagcgcaagcttgagatcatatggcaagaagctctcgagaccacatcgagacatg 417
|||||
QY 61 ttcgacagagtaagtggagacgtgtttccgagaactgtgctggccgacgtgaaggcg 120
|||||
DB 418 ctggagagcgaggagccgtcatctcatctcatctgatatattatcagttcacaaatcact 477
|||||
QY 121 cgggcccctcaacagatcgagagccgacccggaactgtgctgctcgagagccgc 180
|||||
DB 478 aggcagagcttcaatgagatcgagatcccgccacaagaacatcagagctggagccagc 537
|||||
QY 181 atccgagctacagcagctcttctgcaagatggcggtgtgtgtgagagagcgccagc 240
|||||
DB 538 atccgagagctgacagagatggttcatgatatggtcattgttgcagactcagggtgaa 597
|||||
QY 241 accctgaagctacagcagctcttctgcaagatggcggtgtgtgtgagagagcgccag 300
|||||
DB 598 atggtcaacaacatcgagagaaatgtgtgtaactctgtagattacgtggaaactgtgcaag 657
|||||
QY 301 gcgcaggtgcggaagcgctgcaagtaagagagaag 336
|||||
DB 658 gaagagacgaagaagccatcaataaccagagcaag 693
|||||

RESULT 7
AAT62411
ID AAT62411 standard; cDNA; 711 BP.
XX
AC AAT62411;
XX
DT 02-JUL-1997 (first entry)
XX
DE Mouse epimorphin truncated variant 2M coding sequence.
XX
KW Human; mouse; epimorphin; coiled-coil region; functional domain; tissue;
KW hydrophobic; deletion; truncation; regulation; morphogenesis; epithelium;
KW artificial organ; cosmetic; hair tonic; ds.
OS Synthetic.
XX
PN JP09065885-A.
XX
PD 11-MAR-1997.
XX
PR 29-MAR-1996; 96JP-0099684.
XX
PR 19-JUN-1995; 95JP-0175540.
XX
PR 31-MAR-1995; 95JP-0099980.
XX
PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
DR WPI; 1997-220419/20.
XX
DR P-PSDB; 14260.
XX
PT Modified epimorphin and related DNA - useful e.g. for treatment of
XX tissues or in artificial organs, or as an ingredient in cosmetics
XX
PS Example 1; Page 13; 18pp; Japanese.
XX
CC The invention relates to novel human (AAW14257-9) or mouse (AAW14260-2)
CC epimorphin proteins with replacements, deletions or substitutions in the
CC amino acid sequence. The new epimorphin protein consists of: (a) an
CC N-terminal coiled-coil region; (b) a functional domain in the middle;
CC and (c) a C-terminal coiled-coil region. A hydrophobic region in the
CC C-terminal has been deleted and at least some amino acids have been
CC deleted from the terminals of coiled coil regions (a) and/or (c). This
CC sequence encodes the mouse epimorphin protein lacking the N-terminal
CC 29 amino acids. Epimorphin is a protein which regulates morphogenesis

CC of epithelial tissues. It can be used for treatment of tissues or used
CC directly in artificial organs or as an ingredient in cosmetics, hair
CC tonic, etc.
XX
SQ Sequence 711 BP; 232 A; 155 C; 201 G; 123 T; 0 other;

Query Match 9.6%; Score 124.8; DB 18; Length 711;
Best Local Similarity 60.7%; Pred. No. 4.2e-13;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1 atccagcgcaagcttgagatcatatggcaagaagctctcgagaccacatcgagacatg 60
|||||
DB 358 atccagcgcaagcttgagatcatatggcaagaagctctcgagaccacatcgagacatg 417
|||||
QY 61 ttcgacagagtaagtggagacgtgtttccgagaactgtgctggccgacgtgaaggcg 120
|||||
DB 418 ctggagagcgaggagccgtcatctcatctcatctgatatattatcagttcacaaatcact 477
|||||
QY 121 cgggcccctcaacagatcgagagccgacccggaactgtgctgctgagagccgc 180
|||||
DB 478 aggcagagcttcaatgagatcgagatcccgccacaagaacatcagagctggagccagc 537
|||||
QY 181 atccgcaagctacagcagctcttctgcaagatggcggtgtgtgtgagagagcgccag 240
|||||
DB 538 atccgagagctgacagagatggttcatgatatggtcattgttgcagactcagggtgaa 597
|||||
QY 241 accctgaagctacagcagctcttctgcaagatggcggtgtgtgtgagagagcgccag 300
|||||
DB 598 atggtcaacaacatcgagagaaatgtgtgtaactctgtagattacgtggaaactgtgcaag 657
|||||
QY 301 gcgcaggtgcggaagcgctgcaagtaagagagaag 336
|||||
DB 658 gaagagacgaagaagccatcaataaccagagcaag 693
|||||

RESULT 8
AAT16090
ID AAT16090 standard; cDNA; 798 BP.
XX
AC AAT16090;
XX
DT 06-JUN-1996 (first entry)
XX
DE Mouse epimorphin fragment (123).
XX
KW Epimorphin; human; mouse; wound; burn; epithelial tissue;
KW diagnosis; treatment; morphogenetic abnormality; cosmetic;
KW hair growth stimulator; ds.
XX
OS Mus musculus.
XX
PN EP698666-A2.
XX
PD 28-FEB-1996.
XX
PR 20-JUN-1995; 95EP-0304270.
XX
PR 31-MAR-1995; 95JP-0099980.
XX
PR 21-JUN-1994; 94JP-0162874.
XX
PR 31-MAR-1995; 95JP-0099979.
XX
PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
PI Hirai Y, Koshida S, Oka Y;
XX
DR WPI; 1996-118213/13.
XX
PT Novel polypeptide containing an epimorphin functional domain - has
XX possible benefits in epithelial tissue treatments, e.g. burns and
XX for artificial organs
XX
PS Claim 27; Page 51; 62pp; English.

```

XX  WO9308213-A.
PN
XX
XX  29-APR-1993.
PD
XX
XX  15-OCT-1992; 92WO-JP01340.
PF
XX  16-OCT-1991; 91JP-0294856.
PR  16-OCT-1991; 91JP-0294857.
PR  17-APR-1992; 92JP-0122906.
PR  30-APR-1992; 92JP-0135692.
XX
XX  (BIOM-) BIOMATERIAL RES INST CO LTD.
PA
XX  Hirai Y, Takashina M, Takebe K;
PI
XX  WPI; 1993-152423/18.
DR
XX  P-PSDB; AAR36554.
XX
PT  Novel active substance epimorphin, its gene and antibodies - for
PT  diagnosing and treating epithelial diseases
XX
XX  Claim 10; Page 57; 76pp; Japanese.
XX
CC  This sequence encodes one of three isolated isoforms of mouse
CC  epimorphin, a protein produced by mesenchymal cells and which
CC  induces epithelial tissue morphogenesis. Modified forms of
CC  epimorphin in which the C-terminal hydrophobic region is replaced
CC  by a defective or non-hydrophobic peptide are also claimed.
CC  See also AAQ41592 and AAQ41593.
XX
SQ  Sequence 840 BP; 258 A; 193 C; 238 G; 151 T; 0 other;

Query Match
Best Local Similarity 9.6%; Score 124.8; DB 14; Length 840;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

OY  1 atccagcgccagctgtagatcatgagcaagaagctctggcgaccagatcgagacatg 60
DB  445 atccagcgccagctgtagatcatgagcaagaagctctggcgaccagatcgagacatg 504
OY  61 ttcgagcaggtgaagtggagacgtgttttcgagaacttgcgtgcgcgtggaagggcg 120
DB  505 cggagcaggtgaagtggagacgtgttttcgagaacttgcgtgcgcgtggaagggcg 120
OY  121 cggcgccctcaacgagatcgagagcgccacgcggaactgtctgcgcctggagagcg 180
DB  565 aggcgaagctcgaatgagatcgagatcgccgcgaacaaagacatcaatgaagctggagaccagc 624
OY  181 atccgagacgtacacgagctcttcttcgagatgagcggtgtgtgtgagaaagcgccgac 240
DB  625 atccgagagctgtagagatggttcatcagatgagatggttgcgagactcggggtgaa 684
OY  241 accctgaagctatcgagctcaacgtacaaagaagcgtgcactacacgcgcgcgcgcgcgc 300
DB  665 atggtcaacaacaatcgaggaatgtgtgtaactctgttaattacgttggaacatgccaag 744
OY  301 ggcgaggtgcggaagcgctgcagttacgaggaag 336
DB  745 gaagagacgaagaagaagcacaataaccagagcaag 780

```

RESULT 11

```

AAQ75249
ID  AAQ75249 standard; cDNA; 840 BP.
XX
XX  AAQ75249;
AC
XX  10-AUG-1995 (first entry)
DT
XX
DE  Mouse epimorphin isoform B gene.
XX

```

```

KW  Probe; epimorphin; human; mouse; lambda-gt11; expression library;
KW  monoclonal antibody; isoform; drug; congenital; acquired; E.coli;
KW  epidermal abnormality; ds.
XX
XX  Mus musculus.
OS
XX
XX  Key
FH  Location/Qualifiers
FT  variation 790..840
FT  /*tag= a
FT  /note= "sequence variance in isoform B"
XX
XX  JF06293800-A.
PN
XX
XX  21-OCT-1994.
PD
XX
XX  15-OCT-1992; 92JP-0301581.
PF
XX  15-OCT-1992; 92JP-0301581.
PR
XX  15-OCT-1992; 92JP-0301581.
XX
XX  (BIOM-) BIOMATERIAL KENKYUSHO KK.
PA
XX
XX  WPI; 1995-009638/02.
DR  P-PSDB; AAR66481.
DR
XX
XX  Human or murine epimorphin - useful for development of drugs to
XX  treat congenital and acquired epidermal form abnormality
XX
XX  Claim 10; Page 8-9; 41pp; Japanese.
XX
CC  The sequence of the gene encoding mouse epimorphin isoform B. A DNA
CC  fragment (AAQ75250) containing the mouse epimorphin gene (AAQ75247) was
CC  isolated from a lambda-gt11 expression cDNA library screened with a
CC  monoclonal antibody raised against mouse epimorphin. A probe
CC  (AAQ75243) derived from the mouse gene sequence was used to isolate
CC  isoforms of the mouse gene (AAQ75248-9) and the gene encoding human
CC  epimorphin (AAQ75244) and isoforms (AAQ75245-6). The genes were cloned
CC  into expression systems for the production of the protein in E.coli and
CC  in animal cells. The epimorphin can be used in the development of drugs
CC  to treat both congenital and acquired epidermal form abnormality.
XX
SQ  Sequence 840 BP; 258 A; 193 C; 238 G; 151 T; 0 other;

```

```

Query Match
Best Local Similarity 9.6%; Score 124.8; DB 16; Length 840;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

```

```

OY  1 atccagcgccagctgtagatcatggaagaagctctggcgaccagatcgagacatg 60
DB  445 atccagcgccagctgtagatcatggaagaagctctggcgaccagatcgagacatg 504
OY  61 ttcgagcaggtgaagtggagacgtgttttcgagaacttgcgtgcgcgtggaagggcg 120
DB  505 cggagcaggtgaagtggagacgtgttttcgagaacttgcgtgcgcgtggaagggcg 120
OY  121 cggcgccctcaacgagatcgagagcgccacgcggaactgtctgcgcctggagagcg 180
DB  565 aggcgaagctcgaatgagatcgagatcgccgcgaacaaagacatcaatgaagctggagaccagc 624
OY  181 atccgagacgtacacgagctcttcttcgagatgagcggtgtgtgtgagaaagcgccgac 240
DB  625 atccgagagctgtagagatggttcatcagatgagatggttgcgagactcggggtgaa 684
OY  241 accctgaagctatcgagctcaacgtacaaagaagcgtgcactacacgcgcgcgcgcgcgcgc 300
DB  685 atggtcaacaacaatcgaggaatgtgtgtaactctgttaattacgttggaacatgccaag 744
OY  301 ggcgaggtgcggaagcgctgcagttacgaggaag 336
DB  745 gaagagacgaagaagaagcacaataaccagagcaag 780

```

RESULT 12

```

Db      565 aggcgaagctctcaatgagatcgagtcgccccaagaacatcatgaagctggaagaccagc 624
Oy      181 atccgcagctacacagagctctcttcgcagatggcggtgctgtgtggaagcagggccagc 240
Db      625 atccgcagagctgcagcagatggttcatgataatgagcattgttctgcagactcgggtgaa 684
Oy      241 accctgaacgtcatcgagctcagctcaacgtaacaaagacggtcagctacacccggccaagc 300
Db      685 atggtcaacaacatcgaggaatgctgtgaactctgtagattacgtggaacatgccaag 744
Oy      301 ggcgcaggtgcggaagccgtgctgctacgaggaag 336
Db      745 gaagagcgaagaagccatcaataaccaggaag 780

```

RESULT 14

AAO41592 standard; cDNA; 870 BP.

AAO41592;

24-AUG-1993 (first entry)

Mouse epimorphin coding sequence.

vascularisation; induction; epithelial tissue morphogenesis; ds.

Mus musculus.

Location/Qualifiers

1..870

/tag= a

/product= epimorphin

MO9308213-A.

29-APR-1993.

15-OCT-1992; 92WO-JP01340.

16-OCT-1991; 91JP-0294856.

16-OCT-1991; 91JP-0294857.

17-APR-1992; 92JP-0122906.

30-APR-1992; 92JP-0135692.

(BIOM-) BIOMATERIAL RES INST CO LTD.

Hirai Y, Takashina M, Takebe K;

WPI; 1993-152423/18.

P-PSDB; AAR36552.

Novel active substance epimorphin, its gene and antibodies - for

diagnosing and treating epithelial diseases

Claim 10; Page 55; 76pp; Japanese.

This sequence encodes one of three isolated isoforms of mouse

epimorphin, a protein produced by mesenchymal cells and which

induces epithelial tissue morphogenesis. Modified forms of

epimorphin in which the C-terminal hydrophobic region is replaced

by a defective or non-hydrophobic peptide are also claimed.

See also AAO41593 and AAO41594.

Sequence 870 BP; 261 A; 192 C; 253 G; 164 T; 0 other;

```

Db      445 atccagcgcagctggagatcactggagaccacccctgacgacgctggaagagatg 504
Oy      61 ttcgaagagggttaatggaacgtgttctccgaagactgtcggcgaactgtaaggcgcg 120
Db      505 ctggagagcggggaagccgctccatctcatctcgtatattatcagattcaacaatact 564
Oy      121 cgggcgcctcaacagatcgagagcgcgcacccgcaactgtcgccttggagagcgc 180
Db      565 aggcgaagctctcaatgagatcgagtcgccccaagaacatcatgaagctggaagaccagc 624
Oy      181 atccgcagctacacagagctctcttcgcagatggcggtgctgtgtggaagcagggccagc 240
Db      625 atccgaagctgcagcagatggttcatgataatgagcattgttctgcagactcgggtgaa 684
Oy      241 accctgaacgtcatcgagctcagctcaacgtaacaaagacggtcagctacacccggccaagc 300
Db      685 atggtcaacaacatcgaggaatgctgtgaactctgtagattacgtggaacatgccaag 744
Oy      301 ggcgcaggtgcggaagccgtgctgctacgaggaag 336
Db      745 gaagagcgaagaagccatcaataaccaggaag 780

```

RESULT 15

AAO75247 standard; cDNA; 870 BP.

AAO75247;

10-AUG-1995 (first entry)

Wild type mouse epimorphine gene.

Probe: epimorphine; human; mouse; lambda-gt11; expression library;

monoclonal antibody; isoform; drug; congenital; acquired; E.coli;

epidermal abnormality; ds.

Mus musculus.

JP06293800-A.

21-OCT-1994.

15-OCT-1992; 92JP-0301581.

15-OCT-1992; 92JP-0301581.

(BIOM-) BIOMATERIAL KENKYUSHO KK.

WPI; 1995-009638/02.

P-PSDB; AAR66479.

Human or murine epimorphine - useful for development of drugs to

treat congenital and acquired epidermal form abnormality

Claim 10; Page 8; 41pp; Japanese.

The sequence of the wild type gene encoding mouse epimorphine. A DNA

fragment (AAO75250) containing the mouse epimorphine gene was isolated

from a lambda-gt11 expression cDNA library screened with a monoclonal

antibody raised against mouse epimorphine. A probe (AAO75243) derived

from the mouse gene sequence was used to isolate isoforms of the mouse

gene (AAO75248-9) and the gene encoding human epimorphine (AAO75244) and

isoforms (AAO75245-6). The genes were cloned into expression systems for

the production of the protein in E.coli and in animal cells. The

epimorphine can be used in the development of drugs to treat both

congenital and acquired epidermal form abnormality.

Sequence 870 BP; 261 A; 192 C; 253 G; 164 T; 0 other;

Query Match 9.6%; Score 124.8; DB 16; Length 870;

Best Local Similarity 60.7%; Pred. No. 4.3e-13;

Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 14:12:02 ; Search time 66.07 Seconds

(without alignments)
4814.521 Million cell updates/sec

Title: US-09-762-249-13

Perfect score: 1295

Sequence: 1 atccagcgcagctgagat.....aaaaaaaaaaaaaaaa 1295

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq.*

2: /cgn2_6/ptodata/2/lna/5B.COMB.seq.*

3: /cgn2_6/ptodata/2/lna/6A.COMB.seq.*

4: /cgn2_6/ptodata/2/lna/6B.COMB.seq.*

5: /cgn2_6/ptodata/2/lna/PCOTUS.COMB.seq.*

6: /cgn2_6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124.8	9.6	486	3	US-08-493-071-13 Sequence 13, Appl
2	124.8	9.6	564	3	US-08-493-071-12 Sequence 12, Appl
3	124.8	9.6	711	3	US-08-493-071-11 Sequence 11, Appl
4	124.8	9.6	798	3	US-08-493-071-14 Sequence 14, Appl
5	124.8	9.6	840	2	US-08-690-457-14 Sequence 14, Appl
6	124.8	9.6	840	2	US-08-628-187-14 Sequence 14, Appl
7	124.8	9.6	867	1	US-08-690-457-13 Sequence 13, Appl
8	124.8	9.6	867	2	US-08-628-187-13 Sequence 13, Appl
9	124.8	9.6	870	1	US-08-690-457-12 Sequence 12, Appl
10	124.8	9.6	870	2	US-08-628-187-12 Sequence 12, Appl
11	124.8	9.6	2940	1	US-08-690-457-15 Sequence 15, Appl
12	124.8	9.6	2940	2	US-08-628-187-15 Sequence 15, Appl
13	121.6	9.4	911	1	US-08-393-985-5 Sequence 15, Appl
14	117.2	9.1	790	1	US-08-393-985-22 Sequence 22, Appl
15	117.2	9.1	2097	1	US-08-393-985-1 Sequence 1, Appl
16	107.2	8.3	1200	1	US-08-356-397-1 Sequence 1, Appl
17	107.2	8.3	3000	1	US-08-393-985-3 Sequence 3, Appl
18	94.4	7.3	486	3	US-08-493-071-9 Sequence 9, Appl
19	94.4	7.3	564	3	US-08-493-071-8 Sequence 8, Appl
20	94.4	7.3	711	3	US-08-493-071-7 Sequence 7, Appl
21	94.4	7.3	795	3	US-08-493-071-10 Sequence 10, Appl
22	94.4	7.3	834	1	US-08-690-457-8 Sequence 8, Appl
23	94.4	7.3	834	2	US-08-628-187-8 Sequence 8, Appl
24	94.4	7.3	864	1	US-08-690-457-7 Sequence 7, Appl
25	94.4	7.3	864	2	US-08-628-187-7 Sequence 7, Appl
26	94.4	7.3	867	1	US-08-690-457-6 Sequence 6, Appl
27	94.4	7.3	867	2	US-08-628-187-6 Sequence 6, Appl

28	89.6	6.9	1054	1	US-08-393-985-7	Sequence 7, Appl
29	79	5.1	973	1	US-08-393-985-9	Sequence 9, Appl
30	64.8	5.0	7218	1	US-08-232-463-14	Sequence 14, Appl
31	58.4	4.5	10660	2	US-08-267-803B-8	Sequence 8, Appl
32	58.4	4.5	10660	4	US-09-041-886-16	Sequence 16, Appl
33	58	4.5	19124	2	US-08-487-826B-13	Sequence 13, Appl
34	57.8	4.5	5852	1	US-07-867-106-2	Sequence 2, Appl
35	57.6	4.4	1129	4	US-09-227-357-40	Sequence 40, Appl
36	57.2	4.4	2230	4	US-08-378-313-24	Sequence 24, Appl
37	55.2	4.3	2836	3	US-08-747-221B-24	Sequence 24, Appl
38	55.2	4.3	2836	3	US-08-747-221B-26	Sequence 26, Appl
39	55.2	4.3	2836	4	US-09-005-051-24	Sequence 24, Appl
40	55.2	4.3	2836	4	US-09-005-051-26	Sequence 26, Appl
41	55	4.2	1558	1	US-08-455-550-7	Sequence 26, Appl
42	54.4	4.2	1736	3	US-09-182-816-22	Sequence 22, Appl
43	54.4	4.2	1736	3	US-09-182-816-24	Sequence 24, Appl
44	54.4	4.2	1736	3	US-09-471-528-22	Sequence 22, Appl
45	54.4	4.2	1736	3	US-09-471-528-24	Sequence 24, Appl

ALIGNMENTS

```
RESULT 1
US-08-493-071-13
: Sequence 13, Application US/08493071
: Patent No. 6127149
:
: GENERAL INFORMATION:
: APPLICANT: Hirai, Yohei
: APPLICANT: Koshida, Shogo
: APPLICANT: Oka, Yumiko
: TITLE OF INVENTION: MODIFIED EPI MORPHIN
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
: STREET: 99 CANAL CENTER PLAZA, SUITE 300
: CITY: ALEXANDRIA
: STATE: VA
: COUNTRY: USA
: ZIP: 22314
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/493,071
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Price, Robert L.
: REGISTRATION NUMBER: 22,685
: REFERENCE/DOCKET NUMBER: 715-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-684-1111
: TELEFAX: 703-684-1124
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 486 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-493-071-13
:
: Query Match 9.6%; Score 124.8; DB 3; Length 486;
: Best local similarity 60.7%; Pred. No. 5e-19;
: Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
:
: QY 1 atccagcgcagctgagatcatgagcaagaagctctcgagcagcagatcgagacatg 60
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
: Db 133 atccagcgcagctgagatcatgagcaagaagcagcagcagcagcagcagcagc 192
```

Db	418	CTGGAGAGCGGGAGAGCCCTCCATCTTCATCTCGGATATTATATACAGATTCACAAATCACT	477
Qy	121	CGAGCGCGCCCTCAAGAGATCTGAAGCGCGCAACGCGAAGTCGTCGCGCTGGAGAAGCGC	180
Db	478	AGGCAGCTCTCATAGATAGATGATCCCGCCACAAAGCATCATGAAGCTGGAGACCTAGC	537
Qy	181	ATCCGCGAGCTACAGAGATCTCTCTGCAAGATGGCGAGTGCCTGCTGAGAAACGAGCGGAC	240
Db	538	ATCCGAGAGCTGCACAGAGATCTTCATGATATGTGGCCATGTTTGTGCGAGACTCCAGGGTGAA	597
Qy	241	ACCTGAGCGTCATCGAGCTCAAGCTACAAAGACGATCGATACACCGCGAGCGCAAG	300
Db	598	ATGGTCAACACATCGAGAGAGAAATGTTGTTGAACCTCTAGATTACGTGAAACATGCCAAG	657
Qy	301	GCAGAGTGCAGGAAGCGCGTGCAGTACAGAGGAAG	336
Db	658	GAAAGACAGGAAGAACCCATTAATATACAGAGCAAG	693

```

1      RESULT      4
2      US-08-493-071-14
3      ; Sequence 14, Application US/08493071
4      ; Patent No. 6127149
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Hirai, Yohei
7      ; APPLICANT: Koshida, Shogo
8      ; APPLICANT: Oka, Yumiko
9      ; TITLE OF INVENTION: MODIFIED EPIMORPHIN
10     ; NUMBER OF SEQUENCES: 30
11     ; CORRESPONDENCE ADDRESSES:
12     ; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
13     ; STREET: 99 CANAL CENTER PLAZA, SUITE 300
14     ; CITY: ALEXANDRIA
15     ; STATE: VA
16     ; COUNTRY: USA
17     ; ZIP: 22314
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy disk
20     ; COMPUTER: IBM PC compatible
21     ; OPERATING SYSTEM: PC-DOS/MS-DOS
22     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/08/493,071
25     ; FILING DATE:
26     ; CLASSIFICATION: 435
27     ; ATTORNEY/AGENT INFORMATION:
28     ; NAME: Price, Robert L.
29     ; REGISTRATION NUMBER: 22,685
30     ; REFERENCE/DOCKET NUMBER: 715-107
31     ; TELECOMMUNICATION INFORMATION:
32     ; TELEPHONE: 703-684-1111
33     ; TELEFAX: 703-684-1124
34     ; INFORMATION FOR SEQ ID NO: 14:
35     ; SEQUENCE CHARACTERISTICS:
36     ; LENGTH: 798 base pairs
37     ; TYPE: nucleic acid
38     ; STRANDEDNESS: double
39     ; TOPOLOGY: linear
40     ; MOLECULE TYPE: cDNA
41     ; US-08-493-071-14

```

```

Query Match          9.6%; Score 124.8; DB 3; Length 798;
Best Local Similarity 60.7%; Pred. 6.1e-19;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0

OY 1 atccagcgccagcttggagatcatalgtggcagaagatctcggcgaccacagatcagagacatg 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 445 ATCCAGGCCCACTGGAATCATCTGGGAGACCAACCATCATGACGACGACTGGAGAAGATG 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 61 ttccgacgaggtlaagtggagcgtgttttcgcgaactctgtgcgcagcttgaagggcgtg 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

```

RESULT      5
US-08-690-457-14
; Sequence 14, Application US/08690457
; Patent No. 578298
;
GENERAL INFORMATION:
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE
TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANT
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,
ADDRESSSEE: Japan
STREET: 1, Taya-cho
CITY: Yokohama-shi
COUNTRY: Japan
ZIP: 244
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,457
FILING DATE: 16-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,309
FILING DATE: June 15, 1993
APPLICATION NUMBER: 294856/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294857/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 122906/1992
FILING DATE: April 17, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 135692/1992
FILING DATE: April 30, 1992
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 840
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

```

Query Match	9.6%;	Score 124.8;	DB 1;	Length 840;
Best Local Similarity	60.7%;	Pred. No. 6.2e-19;		
Matches 204; Conservative	0;	Mismatches 132;	Indels 0;	Gaps 0;

MOLECULE TYPE: CDNA
US-08-690-457-13

Query Match
Best Local Similarity 60.7%; Pred. No. 6.3e-19;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

9.6%; Score 124.8; DB 1; Length 867;
1 atccagcgccagctgagatcatgcaagaagctcgcggcgagccagcagatgacatg 60
|||||
445 ATCCAGCGCCAGCTGGAGATCACTGGGAGAGACCACTGACGACGAGCTGGAAGATG 504
61 ttcgagcaggtaagtggagcgtgtttccgagaactgctgagcagcgtgaaggcg 120
|||||
505 CTGGAGAGCGGAGAGCGCTGCATCTTCGATATTAATCATGATTAACAATCACT 564
121 cggagcgccctcaagcagatcgagagcgccgacagcgaactgctgcgcctggagagcg 180
|||||
565 AGGCAAGCTCTCAATGATGATCGAGTCCCGCCACAAGACATCATGAGCTGGAGACCA 624
181 atccgagcgtacacgagctcttcttcagatgagcgtgctgtgagagcagcgcgac 240
|||||
625 ATCCGAGACTGCACGAGATGTCATGATATGCGCATGTTTGCAGACTCAGGCTGAA 684
241 accctgaagctcatcgagctcaacgtacaaagcgtgactacaccgcccagcgcaag 300
|||||
685 ATGCTCAACAACATCGAGAGAAATGTGTGTAACCTGTGTAATTAATGGAACATGCCA 744
301 ggcgaggtcggaagcgctgacgtacgagagagaag 336
|||||
745 GAAGAGACGAGAAAGCCATCAATACCAAGCAAG 780

RESULT 8

US-08-628-187-13
Sequence 13, Application US/08628187

Patent No. 5837239

GENERAL INFORMATION:

APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.

TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE DESIGNATED

TITLE OF INVENTION: AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBODIES THEREOF

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA, Japan

STREET: 1, Taya-cho

CITY: Yokohama-shi

COUNTRY: Japan

ZIP: 244

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/628,187

FILING DATE: April 5, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 294856/1991

FILING DATE: October 16, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 122906/1992

FILING DATE: April 17, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 135692/1992

FILING DATE: April 30, 1992

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 867

TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-628-187-13

Query Match
Best Local Similarity 60.7%; Pred. No. 6.3e-19;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

9.6%; Score 124.8; DB 2; Length 867;
1 atccagcgccagctgagatcatgcaagaagctcgcggcgagccagcagatgacatg 60
|||||
445 ATCCAGCGCCAGCTGGAGATCACTGGGAGAGACCACTGACGACGAGCTGGAAGATG 504
61 ttcgagcaggtaagtggagcgtgtttccgagaactgctgagcagcgtgaaggcg 120
|||||
505 CTGGAGAGCGGAGAGCGCTGCATCTTCGATATTAATCATGATTAACAATCACT 564
121 cggagcgccctcaagcagatcgagagcgccgacagcgaactgctgcgcctggagagcg 180
|||||
565 AGGCAAGCTCTCAATGATGATCGAGTCCCGCCACAAGACATCATGAGCTGGAGACCA 624
181 atccgagcgtacacgagctcttcttcagatgagcgtgctgtgagagcagcgcgac 240
|||||
625 ATCCGAGACTGCACGAGATGTCATGATATGCGCATGTTTGCAGACTCAGGCTGAA 684
241 accctgaagctcatcgagctcaacgtacaaagcgtgactacaccgcccagcgcaag 300
|||||
685 ATGCTCAACAACATCGAGAGAAATGTGTGTAACCTGTGTAATTAATGGAACATGCCA 744
301 ggcgaggtcggaagcgctgacgtacgagagagaag 336
|||||
745 GAAGAGACGAGAAAGCCATCAATACCAAGCAAG 780

RESULT 9

US-08-690-457-12
Sequence 12, Application US/08690457

Patent No. 5726298

GENERAL INFORMATION:

APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.

TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE

TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND AN

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,

STREET: 1, Taya-cho

CITY: Yokohama-shi

COUNTRY: Japan

ZIP: 244

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690,457

FILING DATE: 16-AUG-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/078,309

FILING DATE: June 15, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 294856/1991

FILING DATE: October 16, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 294857/1991

FILING DATE: October 16, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 122906/1992

FILING DATE: April 17, 1991

PRIOR APPLICATION DATA:

0/UT 12
 08-628-187-15
 : Sequence 15, Application 'US/08628187
 : Patent No. 5837239
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
 : TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE DESIGNATED
 : TITLE OF INVENTION: AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBODIES THEREO
 : NUMBER OF SEQUENCES: 15
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA, Japan
 : STREET: 1, Taya-cho
 : CITY: Yokohama-shi
 : COUNTRY: Japan
 :
 : ZIP: 244
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:
 :
 : APPLICATION NUMBER: US/08/628,187
 : FILING DATE: April 5, 1996
 :
 : CLASSIFICATION: 435

RESULT 13
US-08-393-985-5
Sequence 5, Application US/08393985
Patent No. 5693476
GENERAL INFORMATION:
APPLICANT: Scheller, Richard H.
TITLE OF INVENTION: Methods and Compositions for Modulation
of Vesicular Release
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Application #1.0, Version #1.25
CURRENT APPLICATION DATA:

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 499)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: www-bio.illn.gov/bdnp/image/image.html
The following repetitive elements were found in this cDNA sequence:
285-313, >AT-rich#low_complexity
Seq primer: M13 Forward.
Location/Qualifiers
1. 499
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3077288"
/clone_1id="NIH_MGC_50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/note="Vector: pTRT3-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 140 a 96 c 62 g 201 t
ORIGIN

Query Match 32.2%; Score 416.4; DB 9; Length 499;
Best Local Similarity 99.5%; Pred. No. 4.3e-36;
Matches 428; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

835 gagcgaaatggccttgggaaacacgcttccttcgcgattccttcacggtctac 894
|||||
7 GAGCGAAATGGCCTTGGGAAAACACGCTCTCTTCGATCTTCATCGGCTCAC 66
|||||
895 ggcctatgcaatccctcccaaatatagatctattctgcctcattccctactattaa 954
|||||
67 -GCTATGCAATTCCTCCCAATATAGATCTTATTTGCTCATTTCCCTACTATTAA 125
|||||
955 aatcaccaacaactctctattctctctccttcactctttaatacttcacag 1014
|||||
126 AATCACCAACAACCTTACTATTCTTCTTCTTCACTTTTAAATATCTTTCAACG 185
|||||
1015 gtaataatttgatatttttccaacaattttaagcactgaataatcgaagaagcactc 1074
|||||
186 GTTATATTTTGGTATTTATTTTCCAAACATTTTAAAGCAGCTGAATATGCAACAGACATC 245
|||||
1075 aaatgaagatcagtcattgttgcgtatcttccgcgtatataaataatttaacattta 1134
|||||
246 AAATGGAAGTATCACTGATGTTTGTGATTTTGTGATTTTGTGATTTTGTGATTTTAA 305
|||||
1135 tattttactgattacatgacatgataatgtaagttaaaataactaataatcactaatt 1194
|||||
306 TATTTTACTTGATTTACATATGACATGATGATGATGATGATGATGATGATGATGATGAT 365
|||||
1195 atatgtacataatgacatgattgtttaaacttttaattgaatgaatgaatgaatgaat 1254
|||||
366 ATATGTACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
|||||

QY 1255 aagacgaanaa 1264
Db 426 AAGACGAACA 435

RESULT 13
LOCUS BF893543
DEFINITION QV1-MT0132-131100-473-906 MT0132 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893543
VERSION BF893543.1 GI:12285106
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 429)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?qt=QV1&t=QV1-MT0132-131100-473-906&t=2000-11-13&t=1>)
Seq primer: puc 18 forward
High quality sequence stop: 429.
Location/Qualifiers
1. 429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1id="MT0132"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 91 a 144 c 135 g 59 t
ORIGIN

Query Match 31.1%; Score 403; DB 10; Length 429;
Best Local Similarity 99.8%; Pred. No. 1.3e-34;
Matches 414; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 99 gctggcgaagcgtgaagggcgcgccgcctcaagagatcgaagccgacgcgga 158
|||||
Db 16 GCTGGCGACGTGAAGGGCGCGCGCCGCCCAACGAGATGAGAGCGCCGACCGCGCA 75
|||||
QY 159 actgctgcgcctggagagcgatcgcgagcgaagcagctctcttcagatggcggt 218
|||||
Db 76 ACTGCTGCCTGGAGAGCGCGATCCGCGACGACGAGCTCTTCTTCTTCTTCTTCTTCTTCTT 135
|||||
QY 219 gctggtggaagcgaagcgacacccctgaacgctcagcgtcaacgtaacgaagacgt 278
|||||
Db 136 GCTGGGGAAGGACGAGCGCGACACCTGAAACGTCATCGAGCTCAACGTAACAAAGACGT 195
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 12:53:01 : Search time 1809.41 Seconds
(without alignments)
9659.810 Million cell updates/sec

Title: US-09-762-249-13

Perfect score: 1295
Sequence: 1 atccagcgcgcgtcgtgagat.....aaaaaaaaaaaaaaaaaaaaa 1295

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_huv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	728	56.2	739	10	B1518895 603061815
2	658.6	50.9	911	10	B1519530 603061815
3	543.2	41.9	571	10	BM149627 TCAP3D91
4	521	40.2	598	10	BF892936 OVI-MT013
5	487.4	37.6	525	10	BF894780 OVI-MT013
6	465.4	35.9	525	10	BF892927 OVI-MT013
7	456.4	35.2	480	10	BF195864 7086C04.x
8	447.8	34.6	470	9	A1275149 q170h12.x
9	443.8	34.0	464	10	BF893539 OVI-MT013
10	429.6	33.2	434	9	AA227632 z98h02.r
11	421.4	32.5	434	9	AA227632 z98h02.r
12	416.4	32.2	499	9	AA227632 z98h02.r
13	403	31.1	429	10	BF893543 OVI-MT013
14	402.4	31.1	426	9	AA262151 z523d09.r
15	397.2	30.7	2219	11	AK017897 Mus muscu
16	393.4	30.4	406	9	AM662909 h183c07.x
17	384	29.7	409	9	AM500048 UI-HF-BNO

C	18	338	26.1	384	9	AA262016
C	19	308.4	23.8	310	9	AM500436
C	20	308	23.8	352	12	AQ988258
C	21	296.6	22.9	361	10	R64498
C	22	294.2	22.7	417	9	AM784298
C	23	269.8	20.8	821	10	BG247715
C	24	235.6	18.2	452	10	R33852
C	25	228.4	17.6	547	12	A2366926
C	26	205.8	15.9	339	10	R33851
C	27	202	15.6	484	9	AM499912
C	28	194.4	15.0	820	10	BF165171
C	29	190.2	14.7	356	10	R65594
C	30	149	11.5	149	9	AA213862
C	31	148.4	11.5	583	10	B1776010
C	32	133.8	10.3	659	10	T25062
C	33	118.4	9.1	166	10	BE798232
C	34	114.4	8.8	812	10	BF784996
C	35	113	8.7	889	11	AK009770
C	36	111.4	8.6	527	9	BE200283
C	37	111	8.6	509	9	BB363966
C	38	109.6	8.5	535	10	BG911547
C	39	108	8.3	451	10	BF041469
C	40	108	8.3	954	10	BE799709
C	41	107.6	8.3	691	10	B1546649
C	42	103	8.0	466	9	AM504759
C	43	101.6	7.8	176	10	BF893531
C	44	101.2	7.8	469	10	BG730112
C	45	100.8	7.8	674	10	BG070771

ALIGNMENTS

RESULT 1
B1518895
LOCUS 603061815F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211074 5',
DEFINITION mRNA sequence.
B1518895
VERSION B1518895.1 GI:15343687
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 739)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM11529 row: n column: 03
High quality sequence stop: 738.
Location/Qualifiers
1..739
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5211074"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6, Site_1: NotI, Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range

```

OY 446 ccaagccgagcctctgcctcgaaggagttgcccacaccttccggaactcatt 505
    |||||||
DB 481 ccaaacccggagactctgcccctcgaaggagattgcccacaccttccggaactcatt 422
OY 506 tagaagaagaacccagcttcaagaatctgcaaacacagcctgtccttggaagatgttag 565
    |||||||
DB 421 tagaagaagaacccagcttcaagaatctgcaaacacagcctgtccttggaagatgttag 362
OY 566 ttgataccgtcgaatcttctcgaagaatagatagattcccaacaagtgtgtcaatgtcat 625
    |||||||
DB 361 ttgataccgtcgaatcttctcgaagaatagatagattcccaacaagtgtgtcaatgtcat 302
OY 626 tataataacccttgacactctaccctcttgacagaagaacgaagaagaactgaag-ctgt 684
    |||||||
DB 301 tatatatacacttctgacactcttaccctcttgacagaagaacgaagaagaactgaag-ctgt 242
OY 685 atctgactga-gggtgaatgtc-tgaagccttcctcctaataaagaactcaagaagaag 742
    |||||||
DB 241 atctgactgaagggtgaatgtc-tgaagccttcctcctaataaagaactcaagaagaag 182
OY 743 tcaattggcactctgctaataagaatcaatcgaatgaatgaacttcaattac--ct 800
    |||||||
DB 181 tcattggggcactctgctaataagaatcaatcgaatgaatgaacttcaattac--ct 122
OY 801 tctccctgaataatctcctgtctgttccatttggagcgaattgg-cttgggaataac 859
    |||||||
DB 121 gtcacatgaaatctcctgtctgttccatttggagcgaattgg-cttgggaataac 62
OY 860 caagcttctccttccgattctcctcgaatcgaatcgaatcgaatcgaatcgaatcgaat 919
    |||||||
DB 61 cagcttctccttccgattctcctcgaatcgaatcgaatcgaatcgaatcgaatcgaat 2
OY 920 a 920
    ||
DB 1 A 1

```

RESULT 3
LOCUS BM149627/c 571 bp mRNA linear EST 30-NOV-2001
DEFINITION TCAP3D9131 Pediatric acute myelogenous leukemia cell (FAB M1)
 Baylor-HGSC project-TCAP Homo sapiens cDNA clone TCAP9131, mRNA sequence.
ACCESSION BM149627
VERSION BM149627.1 GI:17171404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 571)
 Wei, Y., Tsang, Y.T.M., Mel, G., Ku, J.M., Ali-Osman, F.R. Jr.,
 Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.
 Pediatric leukemia cDNA Sequencing Project (2001)
TITLE Unpublished (2001)
JOURNAL Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@ccc.org
 Seq primer: M13 primer.
COMMENT Location/Qualifiers
 1..571
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCAP9131"
 /clone_1="Pediatric acute myelogenous leukemia cell (FAB
 M1) Baylor-HGSC project-TCAP"
 /sex="male"
 /tissue_type="leukophoresis"
 /cell_type="myeloid cell"

```

/def stage="pediatric 6 years"
/lab_host="DH10B"
/notes=vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-Oligo(dT) primer [5'GAGGACCTCGACGGCCGCGAGAG(T)VN
3'; V-A, C, G; N-A, C, G, T] and then dg tailed. Second strand
was primed with a BamHI-dc primer
[5'AGAGAGCTCGGATCCGCGCGCGCAATATATATAT(CT) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda PSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carpinci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y. High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.
DNA Res 4: 1, 61-6, Feb 28, 1997).

```

BASE COUNT 210 a 71 c 105 g 184 t 1 others
ORIGIN

Query Match 41.9%; Score 543.2; DB 10; Length 571;
Best Local Similarity 99.0%; Pred. No. 1e-49;
Matches 567; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

OY 710 ggcctgcctcctaataagaactcaagaagaatcgaatcgaatcgaatcgaatcgaatcgaat 769
    |||||||
DB 571 ggcctgcctcctaataagaactcaagaagaatcgaatcgaatcgaatcgaatcgaatcgaat 512
OY 770 acatgatggaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 829
    |||||||
DB 511 acatgatggaactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 453
OY 830 atttggagcaaatctgagccttgggaacacacacgcttctccttcgaatcgaatcgaatcgaat 889
    |||||||
DB 452 atttggagcgaatctgagccttgggaacacacacgcttctccttcgaatcgaatcgaatcgaat 393
OY 890 tctagcgtatgcaatctcctcccaataatagatcttctcgtcattcctcccttactt 949
    |||||||
DB 392 tctagcgtatgcaatctcctcccaataatagatcttctcgtcattcctcccttactt 334
OY 950 attaaatcacacacacacacacacacacacacacacacacacacacacacacacacacacac 1009
    |||||||
DB 333 attaaatcacacacacacacacacacacacacacacacacacacacacacacacacacacac 274
OY 1010 accaggtataatttggatatttcccaacacacacacacacacacacacacacacacacacacac 1069
    |||||||
DB 273 accaggtataatttggatatttcccaacacacacacacacacacacacacacacacacacacac 214
OY 1070 cactcaaatggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1129
    |||||||
DB 213 cactcaaatggaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 154
OY 1130 attatatttcttgaattatcacaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1189
    |||||||
DB 153 attatatttcttgaattatcacaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 94
OY 1190 ctaatatgtacataatgcaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1249
    |||||||
DB 93 ctaatatgtacataatgcaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 34
OY 1250 attcaagcgaacacacacacacacacacacacacacacacacacacacacacacacacacac 1282
    |||||||
DB 33 atttcaagcgaacacacacacacacacacacacacacacacacacacacacacacacacacac 1

```

RESULT 4
LOCUS BF892936 598 bp mRNA linear EST 18-JAN-2001
DEFINITION OVI-MT01332-081100-443-a09 MT0132 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF892936
VERSION BF892936.1 GI:12284395
KEYWORDS EST.

Query Match 37.6% Score 487.4; DB 10; Length 525;
 Best Local Similarity 99.6%; Pred. No. 1e-43; Indels 1; Gaps 1;
 Matches 499; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

155 gcaactctctgctgagagagccatccgagcagctacagagctctcttcagatg 214
 |||||
 525 GCGAATGCTGCGCTGGAGAGCCGATCCGCGAGTACAGAGCTCTTCTTGCAAGG 466

215 cgtgtcgtgtgagagagcagcagcagcagcagcagcagcagcagcagcagc 274
 |||||
 465 CGGTGCTGGGAG 406

275 cgtgtcagctacacccgagcagcagcagcagcagcagcagcagcagcagcagc 334
 |||||
 405 CGGTGCTGGGAG 346

335 agaaccctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 394
 |||||
 345 AGAACCCTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 286

395 cggcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 454
 |||||
 285 CGGGCGGCGCCAGCCGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227

455 gggcgtcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 514
 |||||
 226 GAGAGCTGTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167

515 aacgcagcttcaagaatctcaaacagcagcagcagcagcagcagcagcagcagcagc 574
 |||||
 166 AACGCAGGTTCAAGAAATGCAACAGCGCTGCTGGAAGAGTGTATATATACG 107

575 tcgagatctctcagtaagaatagatcccaagaatctgcaatctcattatataca 634
 |||||
 106 TCGAGATCTCTCAGTAAGAATAGATCCCAAGAGTGTGCAATGTATATATACG 47

635 cctgcactctacgctctctg 655
 |||||
 46 CCTGCACTCTTACCGTCTAG 26

RESULT 6
 BF892927/c 525 bp mRNA linear EST 18-JAN-2001
 LOCUS BF892927
 DEFINITION OVI-MT0132-081100-443-a05 MT0132 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF892927
 VERSION BF892927.1 GI:12284386
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 525)
 Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 2020263
 CONTACT: Simpson A.J.G.
 LABORATORY of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=OVI-MT0132-081100-443-a05&t3=2000-11-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 5
 High quality sequence stop: 525.
 Location/Qualifiers
 1..525
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MT0132"
 /dev_stage="Adult"
 /note="Organ: marrow; Vector: puc18; Site:1; Sma1; Site:2; Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 95 a 146 c 158 g 126 t
 ORIGIN

Query Match 35.9% Score 465.4; DB 10; Length 525;
 Best Local Similarity 99.0%; Pred. No. 2.3e-41;
 Matches 500; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

152 accgcgaactctgctgagagagcagcagcagcagcagcagcagcagcagcagcagc 211
 |||||
 525 ACCGCAATCTGCTGCGCTGGAGAGCCGATCCGCGAGTACAGAGCTCTTTCGAGA 466

212 tggcgtgtcgtgtgagagagcagcagcagcagcagcagcagcagcagcagcagc 271
 |||||
 465 TGGCGTGTCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406

272 agacggtcagctacacggcagcagcagcagcagcagcagcagcagcagcagcagc 331
 |||||
 405 AGACGCTGACTACACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346

332 agaagaaccctctcgcagacccctctgctctctgctctgctctcctcagtagcagc 391
 |||||
 345 AGAAGAACCCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 286

392 gccgcggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 451
 |||||
 285 GCCGCGGCGCGCCAGCCGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227

452 ccgggagctctgcccctcagcagcagcagcagcagcagcagcagcagcagcagc 511
 |||||
 226 CCGGAGACTCTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167

512 agaagcgcagctcagaatctcagaacacagcagcagcagcagcagcagcagcagcagc 570
 |||||
 166 AGAAGCCGAGTCAAGAAATGCAACAGAGCTGTGCTTGGAAAGATGTTATTAT 107

571 accgtcagatgatctcagtaagaatagatcccaagaatctgcaatctcattatata 630
 |||||
 106 ACCGTGAGATGATCTCAGTAAGAATAGATCCCAAGAGTGTGCA--TCATTATAT 49

631 gacacctgacactcttaccgctctg 655
 |||||
 48 GACACCTTGACACTCTTACCGCTAG 24

RESULT 7
 BF195864/c 480 bp mRNA linear EST 03-NOV-2000
 LOCUS BF195864
 DEFINITION 7086C04.X1 NCI CGAP_K1d11 Homo sapiens cDNA clone IMAGE:3643015 3',
 mRNA sequence.
 ACCESSION BF195864
 VERSION BF195864.1 GI:11083169
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1 (bases 1 to 464)	Dias Neto, E., Garcia Correa, R., Veljovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zagó, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunslein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A.	97 (7), 3491-3496 (2000)	Contact: Simpson A.J.G.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL-MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT 20202653

Contact: Simpson A.J.G.,
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/genhtml2.pl?rl=SV1&t2=QV1-MT0132>)

13100-473-e10&t3=2000-11-13&t4=11

Seq primer: puc 18 forward

High quality sequence stop: 464.

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0132"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site:1: Sma1; Site:2:
Sma2; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      103 a      153 c      143 g      65 t
ORIGIN

```

Query Match 34.0%; Score 439.8; DB 10; Length 464;

Best Local Similarity 98.3%; Pred. NO. 1.4e-38;
Matches 455; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY	87	ttccgaaacttctctgcccacgttgaagggcgcgcgcccctcaacagatctgaag	146
Db	3	TTACGCTAGACGGCTGGCCGACGCTGAAGGGCGGGCGCCCTCAACGAGATCGAGAG	62
QY	147	ccgcgcacgcgaaactgtctgcgcttgagagcgcaatccgcgcgtacagaagctctctt	206
Db	63	CCCGCACCGGGAACTGCTGGCGCTGGAGACCCGATCCGGAGCTACACGAGCTCTTCTY	1222
QY	207	gcagatgcggtactgtgttgagaaagcagggccgcacaccttgaactcatgagctcaagt	266
Db	123	GCAGATGGCGGTGCTGGTGGAGAGACAGCGCCGACACCTTAAGCTCATCGAGCTCAACGT	182
QY	267	acaaaagacgctgcgtactacacacgcgcgcagccaaagcgcaggtgcggaagggcgtgcgta	326
Db	183	ACAAAAGACGCTGAGCTACACCGGCCACGCGCAAGGCGCAGGTGCGGAAGCGCCGTGCAGTA	242
QY	327	cgagggagaaagaaacctgcgcggacacctctgtctctctgtctgtccctgtcctaagtacga	386
Db	243	CGAGGAGAAACCCCTGCGCCGACCCCTCTCTCTCTCTGCTGTGTCCTCGCTCAAGTAGCA	302
QY	387	ggacgcgcgcgcgcgcgcacacgcgccatcccaagacatatgaagcgcgtctgggaagagctcac	446
Db	303	GGCGCGCGCGGGCGCCACCGGCCCATTCAGACATGGAACGGCGCTGGGAAGACG-CAC	361
QY	447	caaaagccggagactctgcctctcaaggaggtgtgcaccaacctttccggaactcagttctt	506
Db	362	CAAAAGCCGGAGCTCTGCGCTCGCAGGGAGGTGCCCAACCTTTCGGAGACTCAGTCTTT	421
QY	507	agaaaagaaacgcacagtttaagaatttgcacaacaactctgc	549
Db	422	AGAAAGAAACGCCAGGTTAAAGATTTCACAAACCGACTTTCG 464	

RESULT	10
LOCUS	AM026277/c
DEFINITION	462 bp mRNA linear EST 27-OCT-1998 w118t05.x1 NCI-CGAP_Brn23 Homo sapiens CDNA clone IMAGE:2529200 3'
ACCESSION	AM026277
VERSION	AM026277
KEYWORDS	EST.
SOURCE	human. Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 462)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

```

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbfp/image/image.html
Seq primer: -400P from G1bco
High quality sequence stop: 442.
Location/Qualifiers
1.462
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2529200"
/clone_1lb="NCI CGAP_Brn23"
/tissue_type="gliblastoma (pooled)"
FEATURES
source

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393.985
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2097 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Rat syntaxin 1A 3' end (encoding amino
INDIVIDUAL ISOLATE: acids 4-288; GenBank M95734)
FEATURE:
NAME/KEY: CDS
LOCATION: 2..859
US-08-393-985-1

Query Match 9.1%; Score 117.2; DB 1; Length 2097;
Best Local Similarity 58.2%; Pred. NO. 4.5e-17;
Matches 206; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy	1	atccagcgccagctgagatcgtggcaaggagtcctcgggcgaccagatcgaggacatg	60
Db	437	ATCCAGAGCGACGTGGAGATCACTGGCCGCGACACACAGTGGAGAGTGGAGACATG	496
Qy	61	ttcagcgagggtaagtggagcgtgtttccgagaaattgtgcccgcacgtgaaggcgcg	120
Db	497	CTGGAGAGTGGGAATCCGCCATCTTGGCTCTGGGATCATGACATCCAGCATCTCG	556
Qy	121	cgggcgccctcaacgagatcgagagccgcccaccggaactgctgcctggagagccgc	180
Db	557	AAGCAGGCCCTCAGTCAGATCGAGACAGGACAGTGGATCATCAAGTTGGAGAACAGC	616
Qy	181	atccgagcgtacacagcgtcttcttgagatggcggtgctggtggaagcagccgac	240
Db	617	ATCCGGGAGCTACAGATATGTTCAATGGACATGGCCATGCTGGTGAGAGCCAGGGGAG	676
Qy	241	accctgaacctcatcgagctcaacgtacaagaagcgtgctgactacaccgagcccaag	300
Db	677	ATGATTGACAGGATCGAGTACAAATGTGGAAACACCGCTGTGGACTACGTGGAGAGGCGGTG	736
Qy	301	gcgcaggtgcggaagccgtgcagtacagaggagagaacccctgccggaccctc	354
Db	737	TCTGACACCAAGAGGCCGCTCAAGTACCAGACGACGACGACGAGGAAGATC	790

Search completed: September 24, 2002, 15:00:52
Job time: 2930 sec

PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

11076861
4 (sites)
THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 2219)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishikawa, K., Nomura, K., Numazaki, R., Ohtsuna, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGACCTCTTTTCTTTTTTTTTCVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGATTAAATAATATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

Location/Qualifiers

1..2219
 `/organism=`"Mus musculus"
 `/strain=`"C57BL/6J"
 `/db_xref=`"MGD:MGI:1907940"
 `/db_xref=`"taxon:10090"
 `/clone=`"5830405C08"
 `/sex=`"male"
 `/tissue_type=`"thymus"
 `/clone_lib=`"RIKEN full-length enriched mouse cDNA library"
 `/dev_stage=`"adult"
224..1087
 `/note=`"data source:SPTR, source key:O75558, evidence:ISS homolog to SYNTAXIN 11 putative"
 `/codon_start=`1
 `/protein_id=`"BAB30994.1"
 `/db_xref=`"GI:12857383"
 `/translation=`"MKDLRAELQLSRSDVQDPDGDNDFADPREDIVETDNILSELS YRVIDIQDENOLLIDLIVRQHLNQVRFTSMRLSSIKRTDTSIAKAIKTREGHITH KRSKMSELGEARHGASHAVARI SHAQYASLAFAQAQAYNOAKMQRNCKIR IQRLEIMGDKMSGEIDMFQGVKWDVFSENLLADLKGAALNEIESRHRELLRG GRIDVHELFLOAMVLVEKDFTLVNIELNVQKTLDTGEAKAQVRKAQVYKKKNPCFV TTCGCCCPGVN"

CDS

polyA_signal
2196..2201
 `/note=`"putative"
polyA_site
2219